

SEQUENCE LISTING

<110> National Institute of Advanced Industrial Science and Technology  
National Institute of Technology and Evaluation  
National Research Institute of Brewing

<120> Pyroglutamyl peptidase and a gene thereof

<130> A21773A

<150> JP 2001-403261

<151> 2001-12-27

<160> 10

<170> PatentIn version 3.1

<210> 1

<211> 2201

<212> DNA

<213> *Aspergillus oryzae*

<220>

<223> Inventor: Machida, Masayuki; Abe, Keietsu; Gomi, Katsuya;  
Inventor: Asai, Kiyoshi; Sano, Motoaki; Kin, Taishin  
Inventor: Nagasaki, Hideki; Hosoyama, Akira; Akita, Osamu  
Inventor: Ogasawara, Naotake; Kuhara, Satoru; Tokunaga, Chikara  
Inventor: Toda, Itaru; Saitoh, Chiaki; Senoh, Akihiro

<220>

<221> source

<222> (1).. (2201)

<223> /organism="Aspergillus oryzae"  
/strain="RIB 40"

<220>

<221> CDS

<222> (1001).. (1111)

<220>

<221> CDS

<222> (1197).. (1901)

<220>

<221> Intron

<222> (1112).. (1196)

<223> /number=1

<400> 1

```
tcaagatctg aatgataaca gatttcctag tccattgtaa cacatcgatg gggttgacgt    60
ggtaatagaa tccaagact gatggaaatc gttctaggtg aagtcaaagt ctagatgatg    120
aataaacaac cagactctag gaaaatgctg gtctagaccc ttgggcgaga aggaatgtgc    180
tgataacgtc tcgctgcctt tcagcggtaa cgctaatact aaagatcaac aaacaatcca    240
ggagcaacca gagcaatcgg tgcgtgttca gtaagtgage agtgagtgca caggagcact    300
cacgtgctaa ccttacaaaa gcagcggcac ccatatcaaa caggaagaag tgggccgtac    360
ggtagttcta ggatgacata ccgaaacccc ttatttgttc gcttaaatag atccctgccc    420
agctttactg atggatttct aatcgcaaag taattgggtg aaataccatc ggtattaacc    480
tagtgaatgg tgattctcaa ccatcgagta caagtcattc tcactattga actttccaaa    540
aagccccgtg aacaagcagt ctgcggtttg ccccggtga agcaaggggg aaattgtcgg    600
tcaggactcg gaacttcgga agcgaagcag aatcggcggg ggccaaaagg catgcgacgt    660
gacagcacct cacatcattc cgggacaata acataggttc aattgcacaa ttgtctcaag    720
aacatgggtg attgtcagat tgatacgtca atcaagcttt gtgggcggtc aagatgaggg    780
gaggtcatgt gccttatcac cttatcgata tcgatatcgc gtgatgccaa gacctgcatg    840
cggtggttgt aatgcggggg aagctccgtc gatatcttga atatatcttt agtccctcct    900
ctctatcctt tttgtggcgt acatagctac cgtgtatata cgaagtaaag gcgttgttcc    960
```

caccactgat tcctagcttg ccttgaccta tccactagcc atg gga gac ttt ggc	1015
Met Gly Asp Phe Gly	
1 5	
cca ccg gtg cca ata ccc gag acg gag gta att ggt ctt gct tcg tca	1063
Pro Pro Val Pro Ile Pro Glu Thr Glu Val Ile Gly Leu Ala Ser Ser	
10 15 20	
tct ttg aca gat cca gaa gag gtc tcg gta ctg gtg aca ggg ttc ggt	1111
Ser Leu Thr Asp Pro Glu Glu Val Ser Val Leu Val Thr Gly Phe Gly	
25 30 35	
gtaagtttat ctgtcttatg cttggcttat gtttctttga ccgtcttggc ttctgattcg	1171
gccctcaaat gctaaaatat actag cca ttc aag acc aac cta gtc aat gcc	1223
Pro Phe Lys Thr Asn Leu Val Asn Ala	
40 45	
tcg tat ttg att gcc tca tct ctg cca gag tcg ctt gac ctt cct tcg	1271
Ser Tyr Leu Ile Ala Ser Ser Leu Pro Glu Ser Leu Asp Leu Pro Ser	
50 55 60	
gcg aag ccg tct gga tcc ggg cct act tct cgt cgg att tca att cat	1319
Ala Lys Pro Ser Gly Ser Gly Pro Thr Ser Arg Arg Ile Ser Ile His	
65 70 75	
gtc cat cca tcg ccc att ccc gtc gct tac tca aca gtg cgg aca act	1367
Val His Pro Ser Pro Ile Pro Val Ala Tyr Ser Thr Val Arg Thr Thr	
80 85 90	
att cca acc atc cta gag gat tac gcc aag tcc cat gga ggt cga cga	1415
Ile Pro Thr Ile Leu Glu Asp Tyr Ala Lys Ser His Gly Gly Arg Arg	
95 100 105 110	
cca gac att gta ctc cat atg gga ata gcg gct aca aga tcg tac tac	1463
Pro Asp Ile Val Leu His Met Gly Ile Ala Ala Thr Arg Ser Tyr Tyr	

115	120	125	
tcg att gag acc aag gcg cat cga gat tct tac cac ttg tcc gat atc			1511
Ser Ile Glu Thr Lys Ala His Arg Asp Ser Tyr His Leu Ser Asp Ile			
130	135	140	
aaa ggc aga atc ggt tat gaa gat ggg gag aag gtt tgg agg gag cag			1559
Lys Gly Arg Ile Gly Tyr Glu Asp Gly Glu Lys Val Trp Arg Glu Gln			
145	150	155	
cag ctc ccg cca gta ctc cag gct ggt cct gcg gcg gat tcc aca gac			1607
Gln Leu Pro Pro Val Leu Gln Ala Gly Pro Ala Ala Asp Ser Thr Asp			
160	165	170	
gta gta cgg aaa gtt ctc cac ccc cag ccg ccc aat gac gac ttt ctc			1655
Val Val Arg Lys Val Leu His Pro Gln Pro Pro Asn Asp Asp Phe Leu			
175	180	185	190
aac acg tgg aag tcg ttt gta tct cct gga gca gac gtc cgg ata tcc			1703
Asn Thr Trp Lys Ser Phe Val Ser Pro Gly Ala Asp Val Arg Ile Ser			
195	200	205	
gag gac gct gga cgc tac ctc tgc gag ttc atc ttt tat aca agt ctg			1751
Glu Asp Ala Gly Arg Tyr Leu Cys Glu Phe Ile Phe Tyr Thr Ser Leu			
210	215	220	
gcc cag gcg ttt caa caa ggc cag cac cga aac gtc gtt ttc ttc cat			1799
Ala Gln Ala Phe Gln Gln Gly Gln His Arg Asn Val Val Phe Phe His			
225	230	235	
gtg cct gga tct tgc gcc gac gag gac atc gag aga ggc acg gat att			1847
Val Pro Gly Ser Cys Ala Asp Glu Asp Ile Glu Arg Gly Thr Asp Ile			
240	245	250	
gca gct gga ttg atc aaa gct ctt gta aga tgt tgg gtt agc gag cag			1895
Ala Ala Gly Leu Ile Lys Ala Leu Val Arg Cys Trp Val Ser Glu Gln			



100	105	110
Ile Val Leu His Met Gly	Ile Ala Ala Thr Arg Ser Tyr	Tyr Ser Ile
115	120	125
Glu Thr Lys Ala His Arg Asp Ser Tyr	His Leu Ser Asp	Ile Lys Gly
130	135	140
Arg Ile Gly Tyr Glu Asp Gly Glu Lys Val	Trp Arg Glu Gln Gln	Leu
145	150	155
Pro Pro Val Leu Gln Ala Gly Pro Ala Ala	Asp Ser Thr Asp	Val Val
165	170	175
Arg Lys Val Leu His Pro Gln Pro Pro	Asn Asp Asp Phe	Leu Asn Thr
180	185	190
Trp Lys Ser Phe Val Ser Pro Gly Ala Asp	Val Arg Ile Ser	Glu Asp
195	200	205
Ala Gly Arg Tyr Leu Cys Glu Phe Ile Phe	Tyr Thr Ser Leu	Ala Gln
210	215	220
Ala Phe Gln Gln Gly Gln His Arg Asn Val	Val Phe Phe His	Val Pro
225	230	235
Gly Ser Cys Ala Asp Glu Asp Ile Glu Arg	Gly Thr Asp Ile	Ala Ala
245	250	255
Gly Leu Ile Lys Ala Leu Val Arg Cys Trp	Val Ser Glu	Gln Val
260	265	270

<210> 3

<211> 1000

<212> DNA

<213> *Aspergillus oryzae*

<220>

```

<221> source
<222> (1)..(1000)
<223> /organism="Aspergillus oryzae"
      /strain="RIB 40"
<400> 3
tcaagatctg aatgataaca gatttcctag tccattgtaa cacatcgatg gggttgacgt    60
ggtaatagaa tccaagact gatggaaatc gttctagggtg aagtcaaagt ctagatgatg    120
aataaacaac cagactctag gaaaatgctg gtctagaccc ttgggcgaga aggaatgtgc    180
tgataacgtc tcgtgcctt tcagcggtaa cgctaatacta aaagatcaac aaacaatcca    240
ggagcaacca gagcaatcgg tgcgtgttca gtaagtgagc agtgagtgca caggagcact    300
cacgtgctaa ccttacaaaa gcagcggcac ccatatcaaa caggaagaag tgggccgtac    360
ggtagttcta ggatgacata ccgaaacccc ttatttggtc gcttaaatag atccctgccc    420
agctttactg atggatttct aatcgcaaag taattgggtg aaataccatc ggtattaacc    480
tagtgaatgg tgattctcaa ccatcgagta caagtcattc tcactattga actttccaaa    540
aagccccgtg aacaagcagt ctgcggtttg ccccggtga agcaaggggg aaattgtcgg    600
tcaggactcg gaacttcgga agcgaagcag aatcggcggt ggccaaaagg catgcgacgt    660
gacagcacct cacatcattc cgggacaata acataggttc aattgcacaa ttgtctcaag    720
aacatgggtg attgtcagat tgatacgtca atcaagcttt gtgggcggtc aagatgaggg    780
gaggtcatgt gccttatcac cttatcgata tcgatatcgc gtgatgcaa gacctgcatg    840
cggtggttgt aatgcggggg aagctccgtc gatatcttga atatatcttt agtccctcct    900
ctctatcctt tttgtggcgt acatagctac cgtgtatata cgaagtaaag gcgttgttcc    960
caccactgat tcctagcttg ccttgaccta tccactagcc                               1000

```

```

<210> 4
<211> 300
<212> DNA
<213> Aspergillus oryzae

```

<220>

<221> source

<222> (1).. (300)

<223> /organism="Aspergillus oryzae"

/strain="RIB 40"

<400> 4

agcggcatgc aggttgctgg tatcgttttg caaagcaaga gcatgggcac tggacgatat 60

atatacttgc atttctatgg cgcggtgcac tatctggggt ccgatgcgc ttttagctgc 120

agtcactcgt gatcatttat ttatagggga cttctgtccc cgtcttttc aggttgagtt 180

atacatgttt cacaggtttt ggatacacta tttaccctct gactactatc gatgaatata 240

gacagttgtc aagcatgata ttgggttcta ccgtattcgt atatgtgtag aatttcccgt 300

<210> 5

<211> 816

<212> DNA

<213> Aspergillus oryzae

<220>

<221> source

<222> (1).. (816)

<223> /organism="Aspergillus oryzae"

/strain="RIB 40"

<220>

<221> CDS

<222> (1).. (816)

<400> 5

atg gga gac ttt ggc cca ccg gtg cca ata ccc gag acg gag gta att 48

Met Gly Asp Phe Gly Pro Pro Val Pro Ile Pro Glu Thr Glu Val Ile



1	5	10	15	
ggt ctt gct tcg tca tct ttg aca gat cca gaa gag gtc tcg gta ctg				96
Gly Leu Ala Ser Ser Ser Leu Thr Asp Pro Glu Glu Val Ser Val Leu				
	20	25	30	
gtg aca ggg ttc ggt cca ttc aag acc aac cta gtc aat gcc tcg tat				144
Val Thr Gly Phe Gly Pro Phe Lys Thr Asn Leu Val Asn Ala Ser Tyr				
	35	40	45	
ttg att gcc tca tct ctg cca gag tcg ctt gac ctt cct tcg gcg aag				192
Leu Ile Ala Ser Ser Leu Pro Glu Ser Leu Asp Leu Pro Ser Ala Lys				
	50	55	60	
ccg tct gga tcc ggg cct act tct cgt cgg att tca att cat gtc cat				240
Pro Ser Gly Ser Gly Pro Thr Ser Arg Arg Ile Ser Ile His Val His				
65		70	75	80
cca tcg ccc att ccc gtc gct tac tca aca gtg cgg aca act att cca				288
Pro Ser Pro Ile Pro Val Ala Tyr Ser Thr Val Arg Thr Thr Ile Pro				
	85	90	95	
acc atc cta gag gat tac gcc aag tcc cat gga ggt cga cga cca gac				336
Thr Ile Leu Glu Asp Tyr Ala Lys Ser His Gly Gly Arg Arg Pro Asp				
	100	105	110	
att gta ctc cat atg gga ata gcg gct aca aga tcg tac tac tcg att				384
Ile Val Leu His Met Gly Ile Ala Ala Thr Arg Ser Tyr Tyr Ser Ile				
	115	120	125	
gag acc aag gcg cat cga gat tct tac cac ttg tcc gat atc aaa ggc				432
Glu Thr Lys Ala His Arg Asp Ser Tyr His Leu Ser Asp Ile Lys Gly				
	130	135	140	
aga atc ggt tat gaa gat ggg gag aag gtt tgg agg gag cag cag ctc				480
Arg Ile Gly Tyr Glu Asp Gly Glu Lys Val Trp Arg Glu Gln Gln Leu				

145	150	155	160	
ccg cca gta ctc cag gct ggt cct gcg gcg gat tcc aca gac gta gta	528			
Pro Pro Val Leu Gln Ala Gly Pro Ala Ala Asp Ser Thr Asp Val Val				
165	170	175		
cgg aaa gtt ctc cac ccc cag ccg ccc aat gac gac ttt ctc aac acg	576			
Arg Lys Val Leu His Pro Gln Pro Pro Asn Asp Asp Phe Leu Asn Thr				
180	185	190		
tgg aag tcg ttt gta tct cct gga gca gac gtc cgg ata tcc gag gac	624			
Trp Lys Ser Phe Val Ser Pro Gly Ala Asp Val Arg Ile Ser Glu Asp				
195	200	205		
gct gga cgc tac ctc tgc gag ttc atc ttt tat aca agt ctg gcc cag	672			
Ala Gly Arg Tyr Leu Cys Glu Phe Ile Phe Tyr Thr Ser Leu Ala Gln				
210	215	220		
gcg ttt caa caa ggc cag cac cga aac gtc gtt ttc ttc cat gtg cct	720			
Ala Phe Gln Gln Gly Gln His Arg Asn Val Val Phe Phe His Val Pro				
225	230	235	240	
gga tct tgc gcc gac gag gac atc gag aga ggc acg gat att gca gct	768			
Gly Ser Cys Ala Asp Glu Asp Ile Glu Arg Gly Thr Asp Ile Ala Ala				
245	250	255		
gga ttg atc aaa gct ctt gta aga tgt tgg gtt agc gag cag gta tag	816			
Gly Leu Ile Lys Ala Leu Val Arg Cys Trp Val Ser Glu Gln Val				
260	265	270		

<210> 6

<211> 16

<212> DNA

<213> Artificial

<220>  
 <223> adaptor  
 <400> 6  
 cgagagcggc cgctac 16

<210> 7  
 <211> 13  
 <212> DNA  
 <213> Artificial  
 <220>  
 <223> adaptor  
 <400> 7  
 gtagcggccg ctc 13

<210> 8  
 <211> 29  
 <212> DNA  
 <213> Artificial  
 <220>  
 <223> primer for amplification of Aspergillus oryzae pyroglutamyl pepti  
 dase cDNA  
 <400> 8  
 gggctgcaga tgggagactt tggcccacc 29

<210> 9  
 <211> 30  
 <212> DNA

<213> Artificial

<220>

<223> primer for amplification of Aspergillus oryzae pyroglutamyl peptidase cDNA

<400> 9

ggggaattcc tatacctgct cgctaacc

30

<210> 10

<211> 312

<212> PRT

<213> Artificial

<220>

<223> amino acid sequence wherein 41 amino acids containing polyhistidine-tag are added at N-terminal of amino acid sequence of SEQ ID NO:2

<400> 10

Met Arg Gly Ser His His His His His His Gly Met Ala Ser Met Thr

1 5 10 15

Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp

20 25 30

Arg Trp Ile Arg Pro Arg Asp Leu Gln Met Gly Asp Phe Gly Pro Pro

35 40 45

Val Pro Ile Pro Glu Thr Glu Val Ile Gly Leu Ala Ser Ser Ser Leu

50 55 60

Thr Asp Pro Glu Glu Val Ser Val Leu Val Thr Gly Phe Gly Pro Phe

65 70 75 80

Lys Thr Asn Leu Val Asn Ala Ser Tyr Leu Ile Ala Ser Ser Leu Pro

	85		90		95
Glu Ser Leu Asp Leu Pro Ser Ala Lys Pro Ser Gly Ser Gly Pro Thr					
	100		105		110
Ser Arg Arg Ile Ser Ile His Val His Pro Ser Pro Ile Pro Val Ala					
	115		120		125
Tyr Ser Thr Val Arg Thr Thr Ile Pro Thr Ile Leu Glu Asp Tyr Ala					
	130		135		140
Lys Ser His Gly Gly Arg Arg Pro Asp Ile Val Leu His Met Gly Ile					
145		150		155	160
Ala Ala Thr Arg Ser Tyr Tyr Ser Ile Glu Thr Lys Ala His Arg Asp					
	165		170		175
Ser Tyr His Leu Ser Asp Ile Lys Gly Arg Ile Gly Tyr Glu Asp Gly					
	180		185		190
Glu Lys Val Trp Arg Glu Gln Gln Leu Pro Pro Val Leu Gln Ala Gly					
	195		200		205
Pro Ala Ala Asp Ser Thr Asp Val Val Arg Lys Val Leu His Pro Gln					
	210		215		220
Pro Pro Asn Asp Asp Phe Leu Asn Thr Trp Lys Ser Phe Val Ser Pro					
225		230		235	240
Gly Ala Asp Val Arg Ile Ser Glu Asp Ala Gly Arg Tyr Leu Cys Glu					
	245		250		255
Phe Ile Phe Tyr Thr Ser Leu Ala Gln Ala Phe Gln Gln Gly Gln His					
	260		265		270
Arg Asn Val Val Phe Phe His Val Pro Gly Ser Cys Ala Asp Glu Asp					
	275		280		285
Ile Glu Arg Gly Thr Asp Ile Ala Ala Gly Leu Ile Lys Ala Leu Val					
	290		295		300

Arg Cys Trp Val Ser Glu Gln Val

305

310